



**You have downloaded a document from
RE-BUŚ
repository of the University of Silesia in Katowice**

Title: Tworzenie i udostępnianie bibliograficznych baz danych na przykładzie bazy CYTBIN

Author: Arkadiusz Pulikowski, Jacek Tomaszczyk

Citation style: Pulikowski Arkadiusz, Tomaszczyk Jacek (2006). Tworzenie i udostępnianie bibliograficznych baz danych na przykładzie bazy CYTBIN. W: D. Pietruch-Reizes, W. Babik (red.), Kierunki i priorytety rozwoju informacji naukowej w kontekście budowania społeczeństwa wiedzy (S. 141-153). Katowice: Polskie Towarzystwo Informacji Naukowej.

© Korzystanie z tego materiału jest możliwe zgodnie z właściwymi przepisami o dozwolonym użytku lub o innych wyjątkach przewidzianych w przepisach prawa, a korzystanie w szerszym zakresie wymaga uzyskania zgody uprawnionego.



UNIWERSYTET ŚLĄSKI
W KATOWICACH



Biblioteka
Uniwersytetu Śląskiego



Ministerstwo Nauki
i Szkolnictwa Wyższego

Arkadiusz Pulikowski
Uniwersytet Śląski, Katowice

Jacek Tomaszczyk
Uniwersytet Śląski, Katowice

Tworzenie i udostępnianie baz danych na przykładzie bazy CYTBIN

Referat omawia problematykę tworzenia i udostępniania baz danych w Internecie na podstawie doświadczeń zdobytych przez autorów w czasie prac nad bibliograficzną bazą danych CYTBIN. Baza ta została zaprojektowana i zaimplementowana z wykorzystaniem programu Microsoft Access. Wyjaśniono przyczyny wyboru tego narzędzia, uwzględniając zarówno jego zalety jak i wady. Wiele miejsca poświęcono opisowi złożonej struktury bazy CYTBIN, wynikającej z natury trudnych w implementacji relacji charakterystycznych dla baz cytowań. Wybór programu MS Access jako narzędzia do projektowania i tworzenia baz danych zdeterminował pośrednio zastosowanie metod publikacji online związanych z tym oprogramowaniem. Najlepszym rozwiązaniem okazała się technika Active Server Pages (ASP). Przedstawiono korzyści i ograniczenia płynące z tego wyboru. Szczegółowo scharakteryzowano bezpłatny serwis IASPHost, umożliwiający udostępnianie baz danych MS Access z wykorzystaniem ASP. Wiele wysiłku twórcy bazy CYTBIN włożyli w przygotowanie kwerend. Opisano problemy z ich tworzeniem zarówno dla wersji offline jak i online bazy.

Creating database and web publishing on the example of CYTBIN database

The paper discusses problems of databases creating and making them available on the Internet and has been based on experience gained by the authors during work on CYTBIN bibliographic database. The base was designed and implemented with the use of Microsoft Access. This tool choice was discussed taking into consideration both its advantages and disadvantages. Much attention was focused on a description of the CYTBIN base complex structure arising of the nature of relations characteristic for citation bases that are difficult to implement. The choice of MS Access as a tool for the databases design and creation indirectly determined the use of online publication methods related to the software. Active Server Pages (ASP) technique appeared the best solution. Benefits and limitations of this choice were presented. A free IASPHost service with access to MS Access databases with the use of ASP was characterised. The authors of CYTBIN base put much effort in queries preparation. Problems with the queries creation both in offline and online base version were discussed.

Ze względu na szybki przyrost informacji systemy informacyjne stają się coraz potężniejsze, gromadząc terabajty danych i udostępniając szereg narzędzi do ich analizy. Podstawowym narzędziem gromadzenia, przetwarzania, wyszukiwania i udostępniania danych są systemy baz danych, będące standardową metodą strukturalizacji zarządzania informacją i stanowiące na ogół część systemu informacyjnego. Dostęp do danych zgromadzonych w bazach realizowany jest na różne sposoby w zależności od przyjętych założeń i ustalonych ograniczeń. Można zauważyć wyraźną tendencję do udostępniania baz danych w Internecie, aby miało do nich dostęp jak najwięcej użytkowników. Wiele serwisów internetowych posiada własne bazy, z których płatnie lub bezpłatnie mogą korzystać użytkownicy na całym świecie. Zastosowanie technologii informacyjnej widać nie tylko w dużych przedsiębiorstwach i organizacjach – coraz częściej osoby indywidualne wykorzystują bazy danych, aby zapanować nad informacją, z którą mają do czynienia na co dzień: PINy, loginy, hasła, adresy i numery telefonów, płytek, biblioteczki itd.

W 2002 roku w Zakładzie Bibliografii i Informacji Naukowej Uniwersytetu Śląskiego zrodził się pomysł stworzenia indeksu cytowań na podstawie polskiego czasopiśmiennictwa z zakresu bibliotekoznawstwa i informacji naukowej. Został powołany kilkusobowy zespół do realizacji tego projektu pod kierownictwem prof. dr hab. Barbary Stefaniak. Początkowo wydawało się, że struktura indeksu cytowań będzie dość prosta i nie przysporzy większych problemów. Jednak w trakcie ustalania założeń wstępnych, kryteriów projektu i zakresu systemu okazało się, że mamy tu do czynienia z niezwykle skomplikowanym zagadnieniem. Problem stanowiły liczne i różnorodne dane bibliograficzne, które miały być przechowywane w bazie. Praca nad projektem i jego implementacją trwała wiele miesięcy, a poprawki i usprawnienia wprowadzane są na bieżąco. W chwili obecnej w bazie CYTBIN dostępne są opisy artykułów wraz z cytowaniami opublikowane w latach 2000-2004, pochodzące z następujących czasopism: „Przegląd Biblioteczny”, „Zagadnienia Informacji Naukowej”, „Praktyka i Teoria Informacji Naukowej i Technicznej”, „Bibliotekarz” i „Roczniki Biblioteczne”. W przygotowaniu biuletyn EBIB.

Po ustaleniu podstawowych wymagań i założeń projektu przystąpiliśmy do wyboru systemu zarządzania bazą danych (SZBD). Nasze poszukiwania rozpoczęliśmy od Mikro CDS/ISIS. Jednak jego możliwości okazały się dalece niewystarczające. ISIS, nie będąc relacyjnym systemem baz danych, przechowuje wszystkie dane w jednej strukturze (nie licząc indeksów). Możliwość stosowania rekordów powtarzalnych oraz podpół rozszerza wprawdzie zastosowanie tego programu, ale w przypadku złożonych systemów ISIS okazuje się niewystarczający. Poza tym korzystanie z podpół jest bardzo uciążliwe zarówno podczas wprowadzania danych, jak i w czasie wyszukiwania, a przecież są to dwie podstawowe operacje dokonywane w systemach baz danych. Chcieliśmy, aby system pozwalał nie tylko na tworzenie kwerend wybierających, ale także udostępniał kwerendy parametryczne oraz był wyposażony w narzędzia statystyczne z możliwościami prezentacji wyników na wykresach różnego typu i publikowania ich na stronach WWW. Od systemu oczekiwaliśmy także możliwości automatyzacji wprowadzania danych, kontroli ich poprawności oraz reagowania na błędy popełnione przez użytkowników, wyświetlając odpowiednie komunikaty. Nie bez znaczenia w systemie jest również możliwość swobodnego projektowania interfejsu użytkownika, z którego będą efektywnie korzystać osoby wprowadzające dane i użytkownicy końcowi. Interfejs graficzny, który jest obecnie dominującą formą interakcji programu komputerowego z użytkownikiem, wykorzystuje wiele elementów usprawniających obsługę systemu, takich jak okna dialogowe, przyciski, listy rozwijane, pola wyboru, których niestety ISIS nie udostępnia. Biorąc pod uwagę wspomniane słabe strony i wady ISISa nie mogliśmy wybrać tego systemu do implementacji bazy CYTBIN, gdyż znacznie ograniczyłoby to możliwości konstrukcyjne i wyszukiwawcze, a co z tym idzie – analizę danych z indeksu cytowań.

Doszliśmy do wniosku, że dobrym narzędziem do realizacji naszego projektu będzie relacyjny systemem zarządzania bazą danych. W relacyjnych bazach danych informacje o obiektach ze świata rzeczywistego są przechowywane w połączonych ze sobą tabelach w postaci wartości atrybutów (cech) opisywanych obiektów, np. *książka* (obiekt), *tytuł* (atrybut), *Pan Wołodyjowski* (wartość atrybutu). Projektując bazę danych należy przede wszystkim określić obiekty i ich atrybuty, czyli ustalić, jakie dane będzie przechowywać baza. Na przykład, jeśli baza będzie gromadzić informacje o użytkownikach biblioteki, utworzymy tabelę *UŻYTKOWNICY* z odpowiednimi atrybutami, opisującymi użytkownika, np. *nazwisko*, *imię*, *data urodzenia* oraz innymi atrybutami odpowiednimi do celów, którym ma służyć baza.

Przy tworzeniu tabel należy pamiętać, że z zgodnie z zasadami projektowania relacyjnych baz danych¹, każda tabela powinna mieć tzw. **klucz główny**. Klucz główny to atrybut lub kombinacja atrybutów, która jednoznacznie określa rekord w tabeli. Klucz główny zawsze jest unikalny i nigdy nie może się powtórzyć w tabeli. Kluczem głównym w tabeli UŻYTKOWNICY może być *nr użytkownika*, ponieważ nie będzie w tabeli dwóch użytkowników o tym samym numerze. Kluczem głównym nie może być natomiast np. *nazwisko*, gdyż możliwe jest, że wśród użytkowników znajdują się osoby noszące to samo nazwisko. Klucze główne są bardzo ważnym elementem struktury bazy, ponieważ służą do łączenia tabel. Rzadko zdarza się, aby baza danych składała się tylko z jednej tabeli. Na ogół bazę stanowi kilka lub kilkanaście połączonych ze sobą tabel, a w przypadku bazy CYTBIN mamy do czynienia z kilkudziesięcioma tabelami. Dlaczego baza składa się z wielu tabel, a nie z jednej? Taka dekompozycja danych minimalizuje redundancję i związane z nią problemy (tzw. anomalie), ułatwia i przyspiesza wprowadzanie danych², umożliwia tworzenie różnych indeksów, zapewnia efektywne i wydajne działanie systemu oraz ogranicza możliwości popełniania błędów, zapewniając w ten sposób spójność bazy.

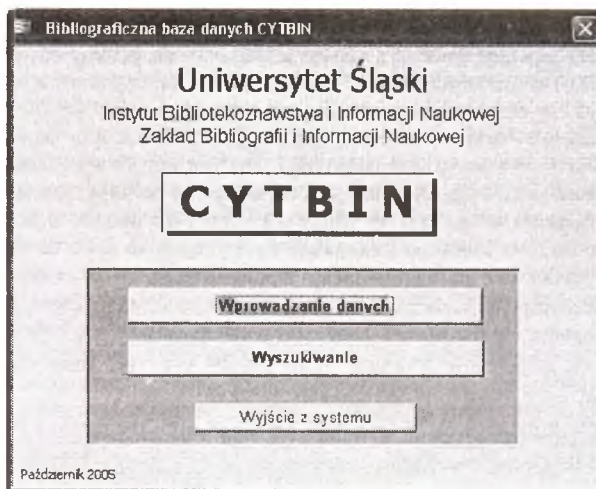
Do relacyjnych systemów baz danych zaliczają się: MS Access, MS SQL, Paradox, My SQL, Informix, Interbase, DB2, Oracle i inne. Do stworzenia indeksu cytowań postanowiliśmy wykorzystać program MS Accessa, który wchodzi w skład pakietu MS Office. Udostępnia on wszystkie niezbędne narzędzia do definiowania danych, zarządzania nimi, wyszukiwania, filtrowania, podsumowania, kontroli poprawności oraz automatyzacji często wykonywanych zadań. Pozwala definiować wiele typów danych (np. tekst, liczba, data, rysunek, dźwięk) i wyświetlać je na różne sposoby. Dzięki regułom poprawności, maskom wprowadzania i własnościom relacji można dodatkowo zabezpieczyć tworzony system przed błędnie wprowadzanymi danymi. Formularze i raporty pozwalają określać sposób prezentacji informacji (w tym również w postaci różnego typu wykresów) oraz wykonywać obliczenia, których wyniki stanowią podstawę statystyk i zestawień. Możliwość korzystania z makroinstrukcji³, języka programowania Visual Basic oraz języka kwerend strukturalnych (SQL) czyni ten program bardzo elastycznym i łatwym w obsłudze narzędziem do tworzenia zarówno najprostszych systemów baz danych, jak i zintegrowanych aplikacji działających w czasie rzeczywistym na setkach milionów rekordów. Dysponuje również dużymi możliwościami w zakresie komunikacji z innymi bazami danych, a także można rozbudować stworzoną w nim bazę tak, aby pracowała np. w systemie MS SQL Server, jednym z najbardziej zaawansowanych systemów baz danych na świecie. Być może fakt, że MS Access wchodzi w skład pakietu biurowego, sprawia, iż jest on niedoceniany i pomijany podczas wyboru SZBD do tworzenia i zarządzania bazami.

Stosując się do zasad projektowania relacyjnych baz danych w systemie CYTBIN utworzyliśmy 18 tabel głównych, by odpowiednio ująć w nich wszystkie potrzebne informacje. Dane zebrano w następujących tabelach: ŹRÓDŁA CYTUJĄCE, AUTORZY ŹRÓDEŁ CYTUJĄCYCH, TYTUŁY CZASOPISM CYTUJĄCYCH, WYDAWNICTWA CIĄGŁE, TYTUŁY WYDAWNICTW CIĄGŁYCH, WYDAWNICTWA ZWARTE, ROZDZIAŁY W KSIĄŻKACH, SZARA LITERATURA, MATERIAŁY KONFERENCJE, INNE DOKUMENTY, TYPY INSTYTUCJI, JEDNOSTKI NADRZĘDNE, JEDNOSTKI PODRZĘDNE, KRAJE, MIASTA, JĘZYKI, SŁOWA KLUCZOWE, ABSTRAKTY. Oprócz tabel głównych baza zawiera kilkanaście tabel łączących. Czym są tabele łączące? Struktura relacyjnej bazy danych, której podstawę stanowią tabele, musi spełniać szereg wymogów. Tabele wchodzące w skład bazy pozostają między sobą w relacjach: jeden-do-jeden, jeden-do-wielu i wiele-do-wielu. Aby wyjaśnić poszczególne typy relacji, posłużymy się przykładami. Tabele ŹRÓDŁA CYTUJĄCE i ABSTRAKTY są ze sobą w relacji jeden-do-jeden, gdyż jednemu rekordowi (dokumentowi) z tabeli ŹRÓDŁA CYTUJĄCE odpowiada **dokładnie jeden** rekord (abstrakt) z tabeli ABSTRAKTY i odwrotnie: jednemu abstraktowi przyporządkowany jest dokładnie jeden dokument. Ten typ relacji bardzo rzadko występuje w bazach danych.

¹ O wszystkich zasadach projektowania relacyjnych baz danych można przeczytać w książce Riordan R.: *Projektowanie systemów relacyjnych baz danych*. Warszawa 2000.

² Na przykład autor książki zostanie wpisany do bazy tylko raz, mimo że napisał więcej niż jedną książkę. Podobnie będzie z dokumentem cytowanym - raz wprowadzony do bazy może być przyporządkowany wielu źródłom cytującym.

³ Makroinstrukcja, zwana w skrócie „makrem”, to grupa czynności wykonywanych w odpowiedzi na zdarzenie, takie jak naciśnięcie przycisku czy ruch myszką.



Rys. 1. Baza CYTBIN w programie MS Access

Tabele JEDNOSTKI NADRZĘDNE i JEDNOSTKI PODRZĘDNE mogą występować w relacji jeden-do-wielu, ponieważ **jednemu rekordowi** z tabeli JEDNOSTKI NADRZĘDNE (np. Uniwersytet Śląski) może być przyporządkowanych **kilka** rekordów z tabeli JEDNOSTKI PODRZĘDNE (np. Biblioteka Uniwersytetu Śląskiego, Instytut Bibliotekoznawstwa i Informacji Naukowej UŚ i wiele innych), ale **jednemu** rekordowi z tabeli JEDNOSTKI PODRZĘDNE przyporządkowany jest **dokładnie jeden** rekord z tabeli JEDNOSTKI NADRZĘDNE (jedyną jednostką nadrzędną dla Biblioteki UŚ jest Uniwersytet Śląski). To najczęściej implementowany typ relacji w relacyjnych bazach danych.

Z powodu długich nazw tabel w bazie CYTBIN do wyjaśnienia relacji wiele-do-wielu posłużymy się innym przykładem. Jeśli chcemy, aby nasza baza przechowywała informacje na temat autorów i ich publikacji, musimy stworzyć dwie tabele główne: AUTORZY i PUBLIKACJE. Te dwie tabele będą w relacji wiele-do-wielu, gdyż jeden autor może napisać wiele publikacji, a jedna publikacja może zostać napisana przez kilku autorów. Mówimy więc, że **jednemu** rekordowi z tabeli AUTORZY może odpowiadać **jeden lub więcej** rekordów z tabeli PUBLIKACJE, a **jednemu** rekordowi z tabeli PUBLIKACJE może odpowiadać **jeden lub więcej** rekordów z tabeli AUTORZY. Ten typ relacji nie może zostać bezpośrednio utworzony w relacyjnym modelu danych. W celu jego zaimplementowania należy posłużyć się *tabelą łączącą*, w której umieszcza się klucze główne obu łączonych tabel, wiążąc w ten sposób autora z jego publikacjami. Osoby, które korzystały z ISISa i jego pól powtarzalnych, często pytają, co zrobić w takim przypadku w Accessie i innych relacyjnych bazach danych, które z założenia nie posiadają pól powtarzalnych. Jedynym rozwiązaniem jest odpowiednie połączenie tabel. Spójrzmy na poniższy rysunek.



Rys. 2. Tabele w relacji wiele-do-wielu

Tabela AUTORZY PUBLIKACJI jest tabelą łączącą. Dzięki niej możemy stworzyć formularz, w którym będzie można przypisać jednej publikacji kilku autorów. W ten sposób otrzymujemy w relacyjnych bazach danych odpowiednik pól powtarzalnych, stosowanych w innych systemach bazodanowych.

Publikacje

Tytuł: Problemy związane z budową bibliograficznej bazy artykułów naukowych

Rok: 2000

Autorzy publikacji:

- Wereszczyńska-Cisło B.
- Siewek R.

Rekord: 14 z 1

Rys. 3. Wprowadzanie informacji o publikacji

Wprowadzanie danych do bazy CYTBIN nie jest skomplikowane. Interfejs graficzny istotnie przyspiesza ten proces, oferując kilka ułatwień. Przede wszystkim są to okna dialogowe, przyciski i pola kombi, które zwalniają użytkownika od wpisywania danych już raz wprowadzonych do bazy oraz nie wymagają od niego znajomości żadnego języka komend.

Źródła cytujące

Rodzaj publikacji: Artykuł 19

Tytuł publikacji: Informacja naukowa o teorii chaosu

Autorzy:

- Janak M., Uniwersytet Jagielloński, Instytut Bibliotekoznawstwa i Informacji Naukowej, Kraków, Polska

Rekord: 14 z 1

Tytuł angielski: Information science and chaos theory

Abstrakt

Czasopismo: PTNT. Praktyka i Teoria Informacji Nauko

Rok: 2000

Tom: 7

Numer/strony: 1

Strony: 4-13

Liczba instytucji: 1

Liczba cytowań: 53

Symbol kłm.

Słowa kluczowe:

- informacja
- teoria chaosu
- informatologia cybernetycz

Wydawnictwa ciągłe | Wydawnictwa zwarte | Rozdziały w książkach | Materiały konferencyjne | Szara literatura | Inne

Dobrychowa M., Informatologia a naukoznawstwo, Aktualne Problemy Informacji i Dokurr

Sordylowa B., Informacja naukowa wśród innych dyscyplin, Przegląd Biblioteczny, 1984,

Sordylowa B., Wzajemne relacje między dyscyplinami: informacja naukowa, bibliologia, b

Sordylowa B., Bibliologia a naukoznawstwo, Przegląd Biblioteczny, 1998, 2/3

Belkin N. J., Progress in documentation: information concepts for information science, Jo

Brookes B. C., The foundation of information science. Part I. Philosophical aspects, Jour

Rekord: 14 z 24

Dodaj

Usuni

Wstaw symbol

ZNAJDŹ

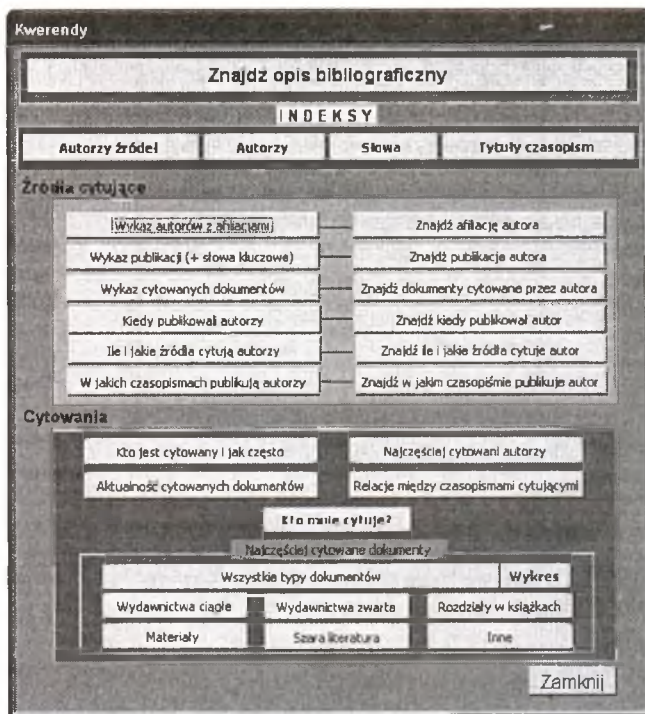
ZAMKNIJ

Pokaż wszystkie rekordy

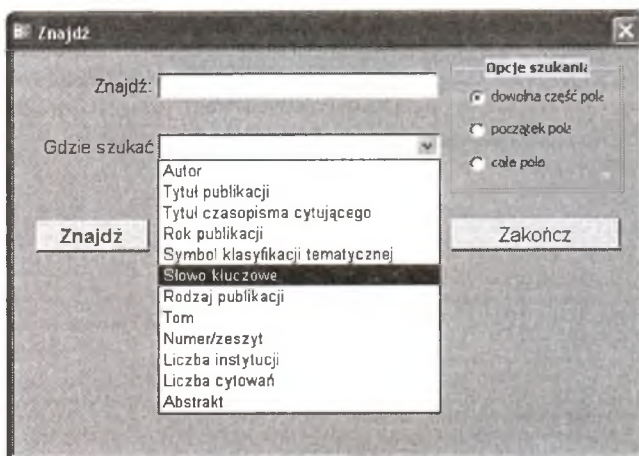
Rekord: 14 z 456

Rys. 4. Wprowadzanie danych do bazy CYTBIN. Górna część okna zawiera informacje o dokumencie cytującym. Dolna część okna to opis dokumentów cytowanych

Wydzielenie w bazie kilkunastu tabel pozwala na tworzenie kwerend, które odpowiedzą niemal na każde pytanie użytkownika, pod warunkiem oczywiście, że potrzebne dane zostały wprowadzone do systemu. Baza zawiera wszystkie niezbędne informacje umożliwiające wieloaspektowe wyszukiwanie bibliograficzne (w tym wyszukiwanie w abstraktach), analizy bibliometryczne oraz wyszukiwanie cytowań.



Rys. 5. Moduł wyszukiwawczy w bazie CYTBIN



Rys. 6. Wyszukiwanie opisów bibliograficznych

Konsekwencją wyboru programu MS Access jako narzędzia do projektowania i tworzenia baz danych było zastosowanie metod publikacji online związanych z tym oprogramowaniem. Do udostępniania w Internecie baz danych utworzonych w MS Access wykorzystuje się technologię ASP stworzoną przez firmę Microsoft. ASP jest akronimem pochodzącym od angielskiego „active server pages”, co można przetłumaczyć jako „aktywne strony serwera”. Aktywne – znaczy tworzone w sposób dynamiczny, strony – to strony WWW, a odniesienie do serwera ma na celu wskazanie, że dynamiczne tworzenie zawartości strony jest wykonywane na komputerze serwera, a nie na komputerze użytkownika. Wyjaśnienia wymaga określenie „dynamiczny”. Stosuje się je w odniesieniu do stron WWW, których zawartość nie jest z góry określona i może się zmieniać zależnie

od wcześniej zdefiniowanych przez twórcę strony warunków⁴ lub podejmowanych przez użytkownika operacji. W obu przypadkach język opisu strony WWW – HTML – jest wzbogacany o dodatkowe instrukcje zapisane w jednym z kilku dostępnych języków skryptowych. To właśnie one są odpowiedzialne za dynamiczną zawartość stron WWW. Języki skryptowe są językami programowania, których polecenia mogą być wykonywane albo przez przeglądarki internetowe po stronie użytkownika albo przez oprogramowanie serwera, na którym strony WWW są zamieszczane. Dobrym przykładem języka pierwszej kategorii jest JavaScript, do drugiej zaś można zaliczyć PHP, CGI, PERL i ASP. Strony WWW zawierające polecenia języka JavaScript są zapisane w plikach o standardowym rozszerzeniu html. W przeglądarce internetowej można w nich podejrzeć elementy tego języka. Inaczej jest w przypadku języków skryptowych drugiej kategorii. Zapisane za ich pomocą instrukcje są wykonywane przez komputer-serwer, a do użytkownika trafia już tylko efekt końcowy w postaci pliku zapisanego w „czystym” HTMLu. Działalność programu jest w tym przypadku zauważalna tylko dzięki rozszerzeniom nazw plików: php, cgi, perl lub asp.

Wśród licznych możliwości technologii ASP w zakresie dynamicznego generowania stron WWW jest również dostęp do baz danych. Wykorzystując tę właściwość ASP, udostępniono bazę CYTBIN w Internecie⁵. Dostęp do bazy odbywa się za pomocą formularza (rys. 7) wykonanego przy użyciu standardowych znaczników języka HTML. Wszystkie interfejsy wyszukiwawcze dostępne z poziomu przeglądarki internetowej oparte są na tego typu formularzach. Większość baz danych zamieszczanych w Internecie może być przez użytkowników jedynie przeszukiwana. Bardzo rzadko dopuszcza się wprowadzanie, modyfikowanie i usuwanie rekordów. Tego typu operacje są zazwyczaj zarezerwowane dla wąskiej grupy osób uprawnionej do dokonywania zmian w bazie. Ponieważ nowe rekordy są wprowadzane do bazy CYTBIN tylko na jednym stanowisku komputerowym, nie było potrzeby tworzenia specjalnego interfejsu umożliwiającego manipulacje na rekordach z poziomu strony WWW. Takie rozwiązanie ma sens tylko przy rozproszonym wprowadzaniu danych i zwykle wiąże się ze spowolnieniem pracy z bazą. Szybciej i prościej wprowadza się nowe rekordy korzystając ze stacji roboczej lub pracując w sieci lokalnej. Po dokonaniu zmian trzeba oczywiście przesłać zmodyfikowaną bazę na serwer.

C Y T B I N

ŹRÓDŁA CYTUJĄCE

Fragment nazwiska:

Fragment tytułu:

Czasopismo:

Rok publikacji:

CYTOWANIA

Fragment nazwiska:

Fragment tytułu:

Rok publikacji:

Typ publikacji:

KTO CYTUJE AUTORA?

Nazwisko:

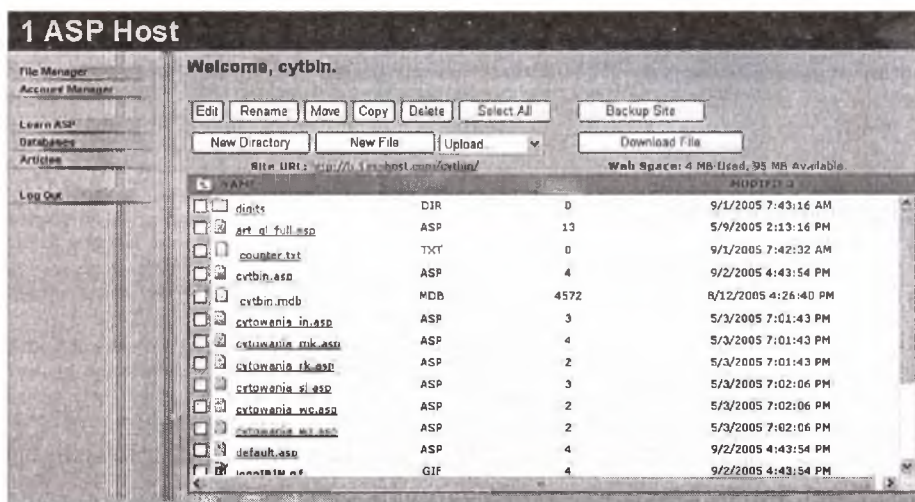
Rys. 7. Formularz dostępny na stronie WWW IBIN UŚ

⁴ Przykładem może być zapisanie przez twórcę strony warunków, które zmieniają rozmiar wyświetlanych elementów w zależności od rozdzielczości ekranu komputera użytkownika.

⁵ Cytbin [dokument elektroniczny]. Tryb dostępu: <http://www.bg.us.edu.pl/bazy/cytbin>

Aby móc korzystać z technologii ASP, potrzebny jest dostęp do serwera z zainstalowanym pakietem Internet Information Services (IIS). Oprogramowanie to jest wbudowane w sieciowe systemy operacyjne firmy Microsoft, takie jak: Windows NT, Windows 2000/2003 Server. Można również doinstalować IIS do systemów Windows 95/98/2000/XP. Wykupienie konta⁶ na serwerze obsługiwany przez np. Windows 2000/2003 Server z możliwością dostępu do baz danych MS Access jest droższe od alternatywnych rozwiązań np. PHP/MySQL⁷. Gdy brak środków finansowych można skorzystać z kont nie wymagających żadnych opłat, pamiętając jednak o ukrytej cenie – pojawiających się od czasu do czasu wyskakujących oknach z reklamami. Przykładem takiego rozwiązania jest 1ASPHost (<http://www.1asphost.com>). Do dyspozycji użytkownika oddano aż 100 MB przestrzeni dyskowej współdzielonej przez pliki ASP, HTML i MDB (bazy danych MS Access). Serwer jest bardzo wydajny. Nawet bardzo skomplikowane kwerendy wykonuje błyskawicznie. Okno z reklamą pojawia tylko przy pierwszym połączeniu ze stroną bazy.

Korzystanie z serwisu 1ASPHost wymaga zarejestrowania – wypełnienia formularza zgłoszeniowego. Użytkownik otrzymuje identyfikator i hasło pozwalające na zalogowanie do systemu. Obsługa konta odbywa się za pomocą wygodnego w użyciu interfejsu dostępnego z poziomu przeglądarki internetowej (rys. 8). Przyciski nawigacyjne pozwalają w intuicyjny sposób przysyłać pliki na serwer, modyfikować je, kasować, a także zapisywać na lokalnym komputerze. Przed wykonaniem jednej z wymienionych operacji należy zaznaczyć pliki, których wybrana czynność ma dotyczyć.



Rys. 8. Interfejs obsługi konta na serwerze 1ASPHost

Kliknięcie nazwy pliku powoduje otwarcie nowego okna przeglądarki i załadowanie do niego strony WWW zapisanej w pliku. Aby przesłać na serwer nowe lub zmodyfikowane pliki należy z rozwijanej listy „Upload...” wybrać „Single Files”. Jednocześnie można przesłać na serwer do 5 plików. Tuż nad panelem wyświetlającym zawartość bieżącego folderu podany jest adres URL tego folderu w Internecie. Nazwa folderu startowego (w przypadku pokazanym na rys. 8 jest to „cytbin”) jest identyczna z nazwą użytkownika (login) podaną w formularzu zgłoszeniowym. Warto dobrze przemyśleć tę nazwę, ponieważ jest ona przypisana na stałe i nie można jej później zmieniać. W celu skrócenia adresu URL przygotowywanej witryny należy plik startowy zapisać pod nazwą default.asp. Plik o tej nazwie jest wczytywany domyślnie, gdy w adresie URL nie sprecyzowano żadnej nazwy pliku. Dzięki temu po wpisaniu w przeglądarce adresu zakończonego samą nazwą folderu ten właśnie

⁶ Konto jest tu rozumiane jako miejsce na dysku serwera, gdzie umieszcza się pliki strony WWW wraz z plikami baz danych.

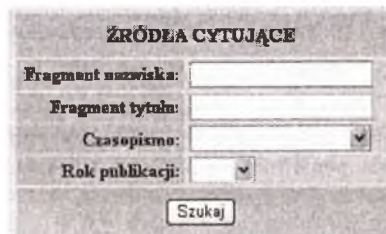
⁷ PHP to skryptowy język programowania służący do tworzenia dynamicznych stron WWW. Pozwala na dostęp do relacyjnych baz danych, wśród których najpopularniejszą jest MySQL.

plik zostanie od razu załadowany⁸. Oprócz plików w formacie asp i mdb na serwerze można również udostępniać pliki o rozszerzeniach nazwy html, htm, css oraz pliki zawierające grafikę. Pamiętając, że pojemność konta to aż 100MB, można z powodzeniem wykorzystywać konto serwisu 1ASPHost do publikowania dużych witryn internetowych.

Ciekawym udogodnieniem oferowanym przez serwis 1ASPHost jest możliwość przesyłania na serwer wielu plików na raz w postaci skompresowanej⁹ (w formacie zip). Oprogramowanie serwera automatycznie rozpakowuje pliki i umieszcza je w wybranym folderze. Takie postępowanie jest szczególnie korzystne przy transferze baz danych. Pliki z bazami MS Access pod wpływem kompresji zmniejszają swoją objętość około pięciokrotnie, co w sposób istotny wpływa na czas ich przesyłania na serwer. Biorąc pod uwagę fakt, że nawet nieduże bazy danych zajmują kilkadziesiąt kilobajtów, jest to istotna oszczędność czasu dla użytkowników 1ASPHost.

Po zapoznaniu się ze środowiskiem pracy, czas na bliższe przyjrzenie się procesowi tworzenia interfejsu pozwalającego na dostęp online do baz danych MS Access. Formularz z rys. 8, jak to już wcześniej zostało napisane, wykonano przy użyciu standardowych znaczników języka HTML. Jest on zapisany w pliku default.asp. Strona z formularzem jest więc startową stroną bazy CYTBIN. Poniżej przedstawiono fragment tego pliku definiujący pierwszy podformularz (rys. 9.), pozwalający na prowadzenie wyszukiwań w obrębie źródeł cytujących, czyli artykułów głównych.

```
<form method="POST" action="szukaj_art_gl.asp">
<table>
<tr><td colspan=2>ŹRÓDŁA CYTUJĄCE
<tr><td>Fragment nazwiska:
  <td><input type="text" name="nazwiskoautora" size="20">
<tr><td>Fragment tytułu:
  <td><input type="text" name="tytulzrodla" size="20">
<tr><td>Czasopismo:
  <td><SELECT name="czasopismo">
    <OPTION>
    <OPTION>Bibliotekarz
    <OPTION>Przegląd Biblioteczny
    <OPTION>PTINT
    <OPTION>ZIN
  </SELECT>
<tr><td>Rok publikacji:
  <td><SELECT name="rok">
    <OPTION>
    <OPTION>2000
    <OPTION>2001
    <OPTION>2002
    <OPTION>2003
  </SELECT>
<tr><td colspan=2><input type="Submit" value="Szukaj" name="button">
</form>
```



Rys. 9. Podformularz odpowiadający podanemu powyżej zapisowi języka HTML

⁸ Dla konta z rysunku 8 będzie to adres: <http://b.1asphost.com/cytbin> Gdyby jako plik startowy miał być użyty plik o nazwie innej niż default.asp wówczas przykładowy adres strony przedstawiałby się następująco: <http://b.1asphost.com/cytbin/startowy.asp>

⁹ W tym celu należy z listy rozwijanej „Upload...” wybrać pozycję „Zip extraction”.

Nie wdając się w szczegóły tworzenia formularzy na stronach WWW¹⁰, można zauważyć, że poszczególne pola tego podformularza zostały umieszczone w oddzielnych komórkach tabeli. Dzięki temu podniesiono przejrzystość. Gdy użytkownik kliknie na przycisku „Szukaj”, zostanie wczytana strona szukaj_art_gl.asp (wskazana w 1 linii kodu). Otwarcie strony będzie towarzyszyć przekazanie wartości atrybutów zdefiniowanych w formularzu. W opisanym przypadku będą to atrybuty: nazwiskoautora, tytułzrodla, czasopismo i rok. Wartości atrybutów będą zależą od tego, co w formularzu wpisał (pierwsze 2 pola) lub wybrał z listy (ostatnie 2 pola) użytkownik.

Plik default.asp nie ma w sobie ani jednego polecenia skryptu ASP. Gdyby nie to, że warto było użyć tej nazwy by skrócić adres URL strony startowej, plik ten mógłby nazywać się jak większość stron startowych – index.html. Skrypty ASP często pojawiają się za to w pliku szukaj_art_gl.asp, który jest wczytywany po kliknięciu przycisku „Szukaj”. Na rysunku 10 pokazano stronę WWW będącą efektem działania skryptów z tego pliku, utworzoną dla trzech kryteriów określonych wcześniej przez użytkownika w podformularzu z rys. 9. Na podstawie danych przekazanych z formularza skrypty ASP przeszukały bazę danych i przedstawiły wynik w formie tabeli.

Wyszukiwany tytuł	Liczba cytowań
1. Babik W.: Język słów kluczowych językoznawstwa sławistycznego - próba charakterystyki Zagadnienia Informacji Naukowej 2000 nr 2 s.74-77,	liczba cytowań: 6
2. Babik W.: Sztuczne sieci neuronowe w systemach informacyjno-wyniskiwawczych Zagadnienia Informacji Naukowej 2000 nr 2 s.43-58,	liczba cytowań: 12
3. Babik W.: Problemy i metody nauki o informacji Zagadnienia Informacji Naukowej 2000 nr 1 s.102-104,	liczba cytowań: 0
4. Babik W.: Zarządzanie informacją we współczesnych systemach informacyjno-wyniskiwawczych - nowe wyzwanie współczesności Zagadnienia Informacji Naukowej 2000 nr 1(75) s.51-63,	liczba cytowań: 11

Rys. 10. Strona WWW wygenerowana przy użyciu skryptów ASP, zawartych w pliku szukaj_art_gl.asp

Skrypty ASP mogą być zapisywane w jednym z czterech języków programowania: VBScript, JScript, PerlScript lub Python¹¹. Większość stron ASP jest tworzonych z wykorzystaniem VBScript. W tym języku wykonane zostały również skrypty obsługujące bazę CYTBIN. Kod źródłowy strony ASP składa się ze standardowych elementów języka HTML rozszerzonych o bloki kodu zapisane w języku skryptowym. Bloki te są wyznaczane za pomocą znaczników <% i %>. Skryptów ASP nie można podejrzewać w przeglądarce internetowej (z menu Widok/Źródło), gdyż wykonuje je serwer. Przykładowo, zamiast:

```
<% Response.write(Request.Form("nazwiskoautora") & "<br>" & Request.Form("czasopismo")) %>
do przeglądarki użytkownika zostanie przekazany tekst: Babik<br>ZIN. Tworzenie stron ASP najlepiej
prześledzić na przykładzie. Plik szukaj_art_gl.asp ma za zadanie pobrać z bazy danych CYTBIN
informacje, o które poprosił użytkownik w formularzu z rys. 9 i wyświetlić wyniki w postaci tabeli,
tak jak to zostało pokazane na rys. 10. Poniżej przedstawiono zawartość tego pliku opatrzoną krótkimi
komentarzami.
```

```
<%@ language="VBScript" codepage=28592%>
```

Pierwsza linia kodu to krótki skrypt ASP definiujący stronę kodową i stosowany język skryptowy. Zapis ten umożliwia poprawne wyświetlanie polskich znaków pojawiających się w tekstach pobieranych z baz danych. Jest konieczny z uwagi na inny sposób kodowania polskich znaków diakrytycznych w bazach MS Access i na stronach WWW.

```
<html>
<head>
<meta name="generator" content="WebSite PRO 4.3" />
<META HTTP-EQUIV="Content-type" CONTENT="text/html; charset=iso-8859-2">
<LINK REL="stylesheet" HREF="style.css" TYPE="text/css">
<title>CYTBIN</title>
</head>
```

¹⁰ Dokładny opis można znaleźć w wielu książkach lub poradnikach internetowych. Dobrym przykładem drugiego z wymienionych źródeł może być kurs Pawła Wimmera dostępny pod adresem: <http://webmaster.helion.pl/kurshtml/form/form.htm>

¹¹ Active Server Pages. W: Wikipedia. Wolna Encyklopedia. http://pl.wikipedia.org/wiki/Active_Server_Pages

Powyższy zapis to typowa część nagłówkowa strony WWW. Oprócz tytułu zawiera ona informacje o nazwie programu, w którym tworzony był dokument, definiuje powszechnie używany standard kodowania polskich znaków oraz wskazuje nazwę pliku, w którym przechowywane są style.

```
<body>
<center>
<table class='glowna'>
<tr><td colspan=2 class='nagl'>
```

Po otwarciu sekcji <body> i wyśrodkowaniu jej zawartości pojawia się zapis tworzący pierwszy wiersz tabeli. Będzie to nagłówek wyświetlający wybrane przez użytkownika kryteria.

```
<%
openStr = "driver={Microsoft Access Driver (*.mdb)};" & _
"dbq=" & Server.MapPath("cytbin.mdb")
Set cn = Server.CreateObject("ADODB.Connection")
cn.Open openStr
```

Rozpoczyna się główna część skryptu. Powyższe wiersze otwierają bazę danych MS Access przechowywaną w pliku o nazwie cytbin.mdb.

```
warunki = ""
ju = ""
au = Request.Form("nazwiskoautora")
ti = Request.Form("tytulzrodla")
j = Request.Form("czasopismo")
yr = Request.Form("rok")
```

Wartości atrybutów pobrane z formularza zostają przypisane zmiennym.

```
If au = "" AND j = "" AND ti = "" AND yr = "" Then
Response.Write ("<tr><td colspan=2 class='nagl'>Nie wybrano żadnych kryteriów")
Else
```

Jeśli wszystkie zmienne są puste, to znaczy, że użytkownik nie wprowadził żadnych kryteriów wyszukiwawczych. Generowany jest komunikat i następuje zakończenie przetwarzania. W przeciwnym wypadku (Else) podejmowane są dalsze czynności.

```
If j = "PTINT" Then ju = "PTINT. Praktyka i Teoria Informacji Naukowej i
Technicznej"
If j = "ZIN" Then ju = "Zagadnienia Informacji Naukowej"
If j = "Bibliotekarz" Then ju = "Bibliotekarz"
If j = "Przegląd Biblioteczny" Then ju = "Przegląd Biblioteczny"
```

Skrócone nazwy czasopism z listy rozwijanej formularza zostają zastąpione pełnymi nazwami.

```
If au <> "" Then warunki = "((Opisy_bibliograficzne.Autor) LIKE ' " & "% " & au &
"%' )"
if ti <> "" AND warunki <> "" Then warunki = warunki & " AND
((Opisy_bibliograficzne.Tytul_publicacji) LIKE ' " & "% " & ti & "%' )"
if ti <> "" AND warunki = "" Then warunki =
"((Opisy_bibliograficzne.Tytul_publicacji) LIKE ' " & "% " & ti & "%' )"
if ju <> "" AND warunki <> "" Then warunki = warunki & " AND
((Opisy_bibliograficzne.Tytul)=" " & ju & " " )"
if ju <> "" AND warunki = "" Then warunki = "((Opisy_bibliograficzne.Tytul)=" " & ju
& " " )"
if yr <> "" AND warunki <> "" Then warunki = warunki & " AND
((Opisy_bibliograficzne.Rok_publicacji)=" " & yr & " " )"
if yr <> "" AND warunki = "" Then warunki =
"((Opisy_bibliograficzne.Rok_publicacji)=" " & yr & " " )"
```

W zależności od wybranych przez użytkownika kryteriów tworzony jest fragment kwerendy, który zostanie użyty poniżej. Jest on przechowywany w zmiennej „warunki”.

```
sql = "SELECT Opisy_bibliograficzne.IDzrodla_cytujacego,
Opisy_bibliograficzne.Autor, Opisy_bibliograficzne.Tytul_publicacji,
Opisy_bibliograficzne.Tytul, Opisy_bibliograficzne.Rok_publicacji,
Opisy_bibliograficzne.Numer_zeszyt, Opisy_bibliograficzne.Strony,
Opisy_bibliograficzne.Liczba_cytowan " & _
```

```
"FROM Opisy_bibliograficzne " & _
"WHERE (" & warunki & ") " & _
"ORDER BY Opisy_bibliograficzne.Autor;"
```

Instrukcja języka SQL sformułowana powyżej pobierze z bazy CYTBIN rekordy odpowiadające kryteriom określonym w zmiennej „warunki”.

```
i=0
Response.Write("Wybrane kryteria:")
If au <> "" Then Response.Write("&nbsp; autor=" & au)
If ti <> "" Then Response.Write("&nbsp; tytuł=" & ti)
If j <> "" Then Response.Write("&nbsp; czasopismo=" & j)
If yr <> "" Then Response.Write("&nbsp; rok=" & yr)
```

W nagłówku tabeli wyświetlone zostają kryteria wybrane przez użytkownika.

```
Set rs = Server.CreateObject("ADODB.Recordset")
rs.Open sql, cn, 3, 3
```

```
On Error Resume Next
rs.MoveFirst
```

Tworzony jest obiekt „rs” przechowujący wynikowy zestaw rekordów. Wskaźnik odczytu/zapisu ustawiany jest na pierwszym rekordzie.

```
Do While Not rs.EOF
i=i+1
Response.Write("<tr><td class='naglewy'>" & i & ".<td><a href=" & Chr(34) &
"art_gl_full.asp?id=" & rs.Fields("IDzrodla_cytujacego") & Chr(34) & ">"_
& rs.Fields("Autor") & ": " & rs.Fields("Tytul_publicacji")_
& "<br>" & rs.Fields("Tytul") & " " & rs.Fields("Rok_publicacji") & " nr " &
rs.Fields("Numer_zeszyt")_
& " s." & rs.Fields("Strony") & ", liczba cytowań: " & rs.Fields("Liczba_cytowan")
& "</a>") & vbCrLf
rs.MoveNext
Loop
```

Zestaw rekordów jest przeglądany jeden po drugim, aż do ostatniego. Dane z odczytywanych rekordów są formatowane do postaci widocznej na rysunku 10. Zmienna „i” jest wykorzystywana do zapisu liczby porządkowej. Opis bibliograficzny artykułów jest wyświetlany w postaci odnośnika. Jego kliknięcie wywoła stronę art_gl_full.asp przechowującą pełną informację o artykule wraz z zamieszczonymi w nim cytowaniami.

```
If i = 0 Then Response.Write("<tr><td colspan=2 class='nagl'>Brak dokumentów
spełniających warunki kwerendy")
```

Gdy nie zostaną znalezione żadne rekordy, wówczas trzeba o tym poinformować użytkownika.

```
rs.Close
Set rs = Nothing
End If
cn.Close
Set cn = Nothing
%>
</table>
</center>
</body>
</html>
```

Na koniec trzeba jeszcze pozamykać otwarte obiekty oraz etykiety języka HTML.

Przedstawiony opis operacji wykonywanych wewnątrz pliku szukaj_art_gl.asp jest oczywiście mocno uproszczony. Jego celem było jedynie przybliżenie zasad tworzenia stron ASP¹². W oparciu o

¹² Więcej informacji na temat tworzenia stron ASP można znaleźć w literaturze fachowej. Godną polecenia jest książka Jima Buyensa pt. „Bazy danych w Internecie”.

te same zasady stworzone zostały wszystkie pozostałe pliki asp zapewniające dostęp do informacji przechowywanych w bazie CYTBIN.

Tworzenie i udostępnianie w Internecie baz danych jest procesem złożonym. Wszystkie jego etapy mają wpływ na uzyskany efekt końcowy. Jeśli zależy nam na stworzeniu dobrego systemu baz danych dostępnego online, to praca nad każdym z etapów musi charakteryzować się należytą starannością. Szczególnie ważna jest faza projektowania. To od niej zależy w największym stopniu to, czy użytkownik będzie w stanie efektywnie wyszukiwać informacje gromadzone w bazie. Błędy popełnione na tym etapie ujawniają się na następnych. Ich poprawa jest pracą i czasochłonna, a często jest wręcz niemożliwa.

Popularny i łatwy w obsłudze program MS Access z pewnością zaspokoi potrzeby wielu projektantów baz danych, a efekty ich pracy można z powodzeniem prezentować w Internecie. Baza CYTBIN jest tego najlepszym dowodem. Mimo złożonej struktury, 466 rekordów źródeł cytujących, powiązanych z 2332 rekordami cytowań można szybko i wygodnie przeszukiwać, a wyniki przeglądać formie przejrzystych tabel. Na podstawie doświadczeń zdobytych w pracy z tą bazą można śmiało stwierdzić, że połączenie MS Access i ASP to efektywny zestaw narzędzi do tworzenia i udostępniania baz danych w Internecie.